Journal of Hydrology 556 (2018) 877–890

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies

Danlu Guo *, Seth Westra, Holger R. Maier

School of Civil, Environmental and Mining Engineering, The University of Adelaide, North Terrace, Adelaide, SA 5000, Australia

ARTICLE INFO

Article history:
Available online 22 March 2016

Keywords:
Scenario-neutral climate impact study
Stochastic generator
WGEN
Exposure space
Inverse approach
Optimization

SUMMARY

Scenario-neutral approaches are being used increasingly for climate impact assessments, as they allow water resource system performance to be evaluated independently of climate change projections. An important element of these approaches is the generation of perturbed series of hydrometeorological variables that form the inputs to hydrologic and water resource assessment models, with most scenario-neutral studies to-date considering only shifts in the average and a limited number of other statistics of each climate variable. In this study, a stochastic generation approach is used to perturb not only the average of the relevant hydrometeorological variables, but also attributes such as the intermittency and extremes. An optimization-based inverse approach is developed to obtain hydrometeorological time series with uniform coverage across the possible ranges of rainfall attributes (referred to as the 'exposure space'). The approach is demonstrated on a widely used rainfall generator, WGEN, for a case study at Adelaide, Australia, and is shown to be capable of producing evenly-distributed samples over the exposure space. The inverse approach expands the applicability of the scenario-neutral approach in evaluating a water resource system's sensitivity to a wider range of plausible climate change scenarios.
© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Scenario-neutral approaches are being used increasingly to assess the possible impact of climate change on the performance of water resources systems (Brown et al., 2012; Brown and Wilby, 2012; Dessai and Hulme, 2004; Nazemi and Wheater, 2014), as well as social and ecological systems (Gao et al., 2016; Poff et al., 2015). The information generated from these approaches can be used to assess system vulnerability under alternative climate change scenarios, and to calculate climatic thresholds at which system performance begins to change abruptly (Brown et al., 2011; Poff et al., 2015). Scenario-neutral approaches can also accommodate changes in climate projections without the need for additional analysis (Prudhomme et al., 2010), and can help to identify the important hydrometeorological variables, or particularly critical states of these variables that affect the system under consideration. The latter feature is particularly useful for selecting: (1) climate models; (2) strategies to generate future rainfall conditions from GCM-based projections (known as statistical downscaling); and (3) alternative 'lines of evidence' (e.g. expert opinion and data from the paleo-climatic record) that can provide useful information about these variables. Ultimately, this allows for the development of a more complete set of projections that describe how these variables might change in a greenhouse gas-enhanced climate (Nazemi et al., 2013; Singh et al., 2014; Steinschneider and Brown, 2013; Vano et al., 2015).

Central to the scenario-neutral approach is the analysis of system sensitivity to a range of hydrometeorological conditions. Such analyses involve exposing the system to perturbed hydrometeorological forcing data that reflect various hydrometeorological conditions that the system may confront in the future (referred to as the 'exposure space'). To this end, it is important to consider the possible variations not only in the average states of the relevant hydrometeorological variables, such as annual average rainfall and potential evapotranspiration (see Kay et al., 2014; Prudhomme et al., 2013), but also their other attributes, including extremes, seasonality and interannual variability (Meselhe et al., 2009; Moody and Brown, 2013; Prudhomme et al., 2010; Steinschneider and Brown, 2013). Indeed, assessments of historical and/or future changes to rainfall as a result of climate change have already indicated different changes to the averages (Collins et al., 2013), extremes (Ajami et al., 2007; Alexander et al., 2006; Westra et al., 2013, 2014), temporal distribution (Rajah et al., 2014) and low-frequency variability (e.g. Johnson et al., 2011) of rainfall throughout the world. Similarly complex changes to other relevant hydrometeorological variables might also be expected, including potential evapotranspiration, and snowfall and melt.

* Corresponding author.
E-mail address: Danlu.Guo@Adelaide.edu.au (D. Guo).

http://dx.doi.org/10.1016/j.jhydrol.2016.03.025
0022-1694/© 2016 Elsevier B.V. All rights reserved.

One approach to generating perturbed hydrometeorological forcing data is by applying scaling factors to historical records of each of the relevant hydrometeorological variables. These factors can be applied at annual or monthly scales (Kay et al., 2014; Paton et al., 2013; Prudhomme et al., 2013, 2010; Singh et al., 2014), or different factors that can be applied across different quantiles in the entire distribution (Nazemi et al., 2013). Although such approaches might be viable for perturbing a small number of hydrometeorological variables and their attributes (i.e. low-dimensional exposure spaces), the capacity of these to represent the potentially complex variations in a wider range of variables and attributes (i.e. high-dimensional exposure spaces) is likely to be limited. Consequently, when using scaling factors to perturb historical data for climate impact assessments, the resultant projections may not show the full range of variability that can be expected in a greenhouse gas-enhanced climate (Prudhomme et al., 2013, 2010; Steinschneider and Brown, 2013).

The use of stochastic generators has been proposed as an alternative to scaling factors to generate hydrometeorological data in a way that can account for a wider range of possible changes (Whateley et al., 2014). Some recent advances include the use of a multi-site weather generator that is capable of producing realistic time series of meteorological variables with shifts to the mean, standard deviation, extremes, daily-scale Markov transition probabilities and low-frequency (interannual) variability (for examples see Steinschneider and Brown, 2013; Wilby et al., 2014; Yates et al., 2015). This is achieved through the perturbation of stochastic model parameters (including the transition probabilities and the autocorrelation coefficient) and the subsequent application of quantile correction, which, in combination, can be used to generate the high-dimensional exposure space. A challenge with this approach, however, is that it is difficult to assess *a priori* which parameters of the stochastic generator should be modified to produce time series at pre-specified points in the exposure space, potentially leading to insufficient exploration of the exposure space. This challenge arises both as a result of the non-linear mapping between the parameters of a stochastic generator and the statistics of the hydrometeorological variables, as well as due to the stochastic nature of the model, which means that a single parameter set will produce hydrometeorological data that span multiple points on the exposure space (Steinschneider and Brown, 2013).

In order to address the shortcomings of existing approaches in generating hydrometeorological data to form the exposure space, we introduce the concept and framework for an inverse approach with demonstration on a case study. The proposed inverse approach enables stochastic generators to be used to generate time series that uniformly span the desired range of the hydrometeorological variables and attributes of interest, and thus provides uniform coverage of the exposure space to serve the needs of scenario-neutral climate impact assessments. Although generally applicable to any parametric weather generator, this paper focuses on applying the method to rainfall time series for the following reasons:

1. Although stochastic generators have been used to generate a range of weather variables, including temperature, humidity, and wind (e.g. Racsko et al., 1991; Semenov and Brooks, 1999), the majority of applications have focused on the generation of rainfall data, due to their importance as an input to many water resource assessments (e.g. Chiew and McMahon, 2002b; Jones and Thornton, 1993).
2. At daily or shorter timescales, rainfall is intermittent, highly skewed (with rainfall series typically exhibiting a large number of moderate rainfall days and a small number of very heavy rainfall days), and exhibits variability at seasonal, interannual and longer time scales (Bastola et al., 2011; Dubrovský et al., 2000). As a result, rainfall is often regarded as a particularly challenging variable to simulate stochastically.
3. There has been a substantial amount of work on developing stochastic generation models to both generate replicates of historical rainfall data (Beven, 1987; Boughton and Droop, 2003; Chen and Brissette, 2014; Clark and Slater, 2006; Frost, 2004; Furrer and Katz, 2008; Langousis and Kaleris, 2014; Langousis et al., 2015), as well as downscaling GCM-based climate projections (Allen and Pruitt, 1986; Bastola et al., 2011; Fowler et al., 2007; Jones et al., 2011; Kay and Jones, 2012; Wilby et al., 2014; Yates et al., 2015).

The remainder of this paper is structured as follows. In Section 2, we illustrate the alternative approaches that are currently available for generating an exposure space, including the historical scaling, forward and inverse approaches. This section also provides details of the proposed inverse approach. Section 3 introduces a case study and two stochastic generators that are used to illustrate both the proposed approach, as well as a simple forward approach that is used as a basis of comparison. The results are given in Section 4, followed by conclusions in Section 5.

2. Proposed inverse approach to exposure space generation

2.1. Rationale for an inverse approach to perturbing stochastic model parameters

As described in the introduction, a central feature of scenario-neutral approaches is the exploration of a water resource system's response to a range of different hydrometeorological conditions. This range of hydrological variables (e.g. rainfall, temperature, evapotranspiration) and the set of attributes of these variables (e.g. annual average, variance, seasonal differences, extremes) are collectively referred to as an 'exposure space', and represent the range of conditions of interest that a system may be exposed to under a future climate. For example, if a scenario-neutral approach was to be used to evaluate system sensitivity to changes in the average, variability and extremes of rainfall, then this would require generating a three-dimensional exposure space with each dimension representing one of the rainfall attributes.

Fig. 1 illustrates the conceptual approaches that could be used to generate an exposure space **E**, which consists of the plausible future changes (represented as the gray shaded region, with the origin corresponding to no change) in various rainfall attributes of interest (represented by two axes A_1 and A_2, which refer to two generic rainfall attributes or groups of attributes). Two techniques are involved in the perturbation approaches – namely scaling of rainfall time series and stochastic rainfall generation (as shown in the two green squares). We use the term 'scaling' in the figure to collectively refer to perturbations that are directly applied to rainfall time series, through the use of change factors at annual, monthly or other time scales (Kay et al., 2014; Prudhomme et al., 2010; Prudhomme and Williamson, 2013), or more complex methods, such as quantile mapping (as used in Steinschneider and Brown, 2013). Consequently, the scaling technique can only modify rainfall intensity on wet days. The term 'stochastic generation' in the figure refers to indirect modification of the rainfall time series through changing the parameters of stochastic generators (as used in Dubrovský et al., 2000; Jones and Page, 2001; Steinschneider and Brown, 2013). The parameter space Θ consists of two axes of θ_1 and θ_2, which refer to two generic parameters or groups of parameters. The plausible ranges for all parameters are represented by the gray shaded region, while the

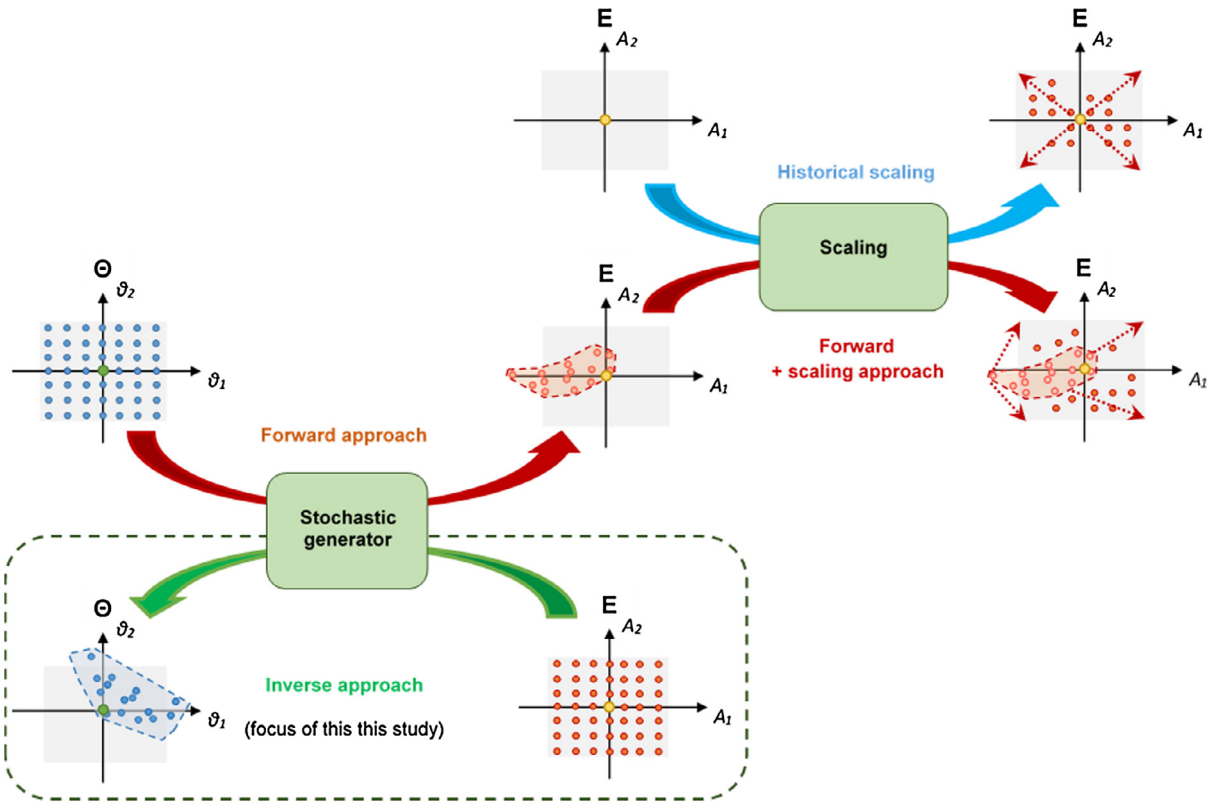

Fig. 1. Outline of alternative approaches to developing an exposure space **E** for scenario-neutral climate impact assessments. Refer to main text for an explanation of each figure element. The focus of this study is on the approaches enclosed in the black dashed box.

origin represents the set of parameters corresponding to the historical rainfall condition.

The first approach, 'historical scaling' (as shown in the top-right corner of Fig. 1) is analogous to the approach used by Prudhomme et al. (2010, 2013), and Kay et al. (2014), in which additive and/or multiplicative scaling factors are applied directly to historical hydrometeorological time series to obtain the desired changes in the relevant variables (usually rainfall and potential evapotranspiration). Although conceptually simple, this approach is not capable of representing variations in the rainfall intermittency, such as the frequency and persistence of dry-/wet-day occurrence. Furthermore, it is difficult to apply this approach to higher-dimensional exposure spaces, since it becomes difficult to develop an approach to scale each attribute independently of the other attributes. Consequently, it can be difficult to sample some regions of the exposure space.

The remaining approaches use stochastic weather generators to obtain perturbed rainfall time series. The 'forward' approach (as illustrated in the middle of Fig. 1) involves perturbing the parameters of stochastic generators over some pre-defined 'parameter space' to yield an exposure space (Dubrovský et al., 2000; Jones and Page, 2001). However, the non-linear mapping between the parameters of a stochastic generator and the attributes of the hydrometeorological variables means that it is unlikely that the full range of the desired exposure space will be covered. Conversely, some perturbations may lead to rainfall attributes with levels out of the defined plausible ranges of the exposure space. Consequently, further scaling may still be necessary after application of the forward approach. Steinschneider and Brown (2013) used this combined 'forward-plus-scaling' approach by firstly perturbing the parameters of a stochastic generator (including Markov chain transition probabilities and the autoregressive model for low-frequency variability) to obtain stochastic sequences without changing the historical rainfall intensity; the wet-day rainfall intensity in the stochastic sequences was subsequently quantile-mapped to yield a set of target daily rainfall series with desired levels of rainfall attributes. Although this approach is likely to provide a much better coverage of the exposure space, some portions of the exposure space may still remain poorly represented because of the difficulty in finding parameters that will result in all combinations of the hydrometeorological attributes of interest.

The limitations of both the historical scaling and forward approaches motivate the 'inverse' approach proposed in this paper (bottom of Fig. 1). Here, the desired values of the attributes of interest in the exposure space are the starting point for the analysis, followed by an optimization step to identify the stochastic generator parameters that produce stochastic sequences with these attributes. This approach provides control over the level of coverage of the exposure space, as required for the implementation of scenario-neutral approaches to climate impact assessments.

2.2. Overview of the inverse approach

To generate hydrometeorological time series with a range plausible attribute levels, the inverse approach is proposed as follows, which involves two primary steps:

(1) *Identify a set of 'target' levels for each attribute included in the exposure space.* In order to achieve an even coverage of the exposure space, we first select the desired levels we would like to sample for each attribute included in the exposure space (referred to as 'target levels'). A number of different approaches can be used to select and combine the target levels (which produce individual 'target locations' in the exposure space), including sampling on a regular grid, or using more computationally efficient sampling methods, such as Latin hypercube sampling (Stein, 1987) or Hammersley sampling (Halton, 1960; Hammersley, 1960).

(2) *Generate hydrometeorological time series that satisfy each target set of attributes.* For each target location of the exposure space, we combine stochastic weather generation with a formal optimization approach to identify the best-fit parameter set for the stochastic generator. This parameter set should be capable of producing hydrometeorological time series with the levels of attributes corresponding to that particular target location, as detailed below.

During the optimization process, the decision variables are the parameters of the stochastic generator. The objective is to identify the parameters of the stochastic rainfall generator that minimize the difference between the values of the hydrometeorological attributes that correspond to the target location and those of the corresponding simulated values. The following objective function is proposed for minimization:

$$F_{obj,i} = \sqrt{\sum_{k=1}^{K} \left\{ \left[\frac{(P_{s,i}^k - P_{his}^k)}{P_{his}^k} - \frac{(\widehat{P_i^k} - P_{his}^k)}{P_{his}^k} \right] * 100\% \right\}^2} \quad (1)$$

where $i = 1, 2, \ldots n$ for n target locations in the exposure space. For the kth attribute of the hydrometeorological variable of interest (P^k), P_s^k represents the target level and $\widehat{P}^k$ represents the simulated level from the stochastic generator. Since different attributes are likely to consist of different magnitudes, the difference between a target level and the simulated level is represented as a percentage change relative to its long-term averaged historical value (P_{his}^k) to ensure consistent scales across different attributes. The optimization problem can be solved using a variety of optimization algorithms, such as genetic algorithms (Holland, 1975) or shuffled complex evolution (Duan et al., 1993).

The optimization procedure proceeds as follows: (1) values of the parameters of the stochastic generator are perturbed based on the searching strategy of the selected optimization algorithm; (2) the corresponding time series of the desired hydrometeorological variables are generated; (3) the values of the attributes of interest are calculated; and (4) the objective function is calculated in accordance with Eq. (1), which drives the algorithm's searching behavior. Steps (1)–(4) are then repeated until the specified stopping criterion has been met, such as the completion of a prespecified number of iterations or until the objective function value is sufficiently small (Maier et al., 2014). It is important to note that as part of the inverse approach, the random seed of the stochastic generator should be held constant to ensure that the optimization proceeds as efficiently as possible, as discussed further in the following section.

2.3. Random sampling issues of stochastic generators and the implications on the inverse approach

The stochastic component of the rainfall generator can produce substantial variations in the simulation of rainfall attributes, even with a single parameter set. This randomness can affect the efficiency of the optimization process used in the inverse approach. Essentially, every iteration of the optimization involves a comparison among multiple parameter sets in terms of their ability to generate the target locations in the exposure space. However, as a result of stochastic generation, a single parameter set can lead to multiple potential locations on the exposure space (Fig. 2). This can then mislead the comparison and affect optimization efficiency, as changes made to parameters by the optimization algorithm in order to lead the search in one particular direction might actually have the opposite effect.

To illustrate this issue, consider a simple optimization problem to find the best-fit parameters of a Gaussian distribution with the

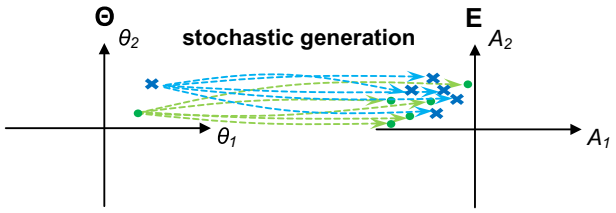

Fig. 2. Using a single parameter set in the parameter space Θ can yield multiple points (shown by the dashed arrows) in the exposure space **E** as result of the stochastic nature of the weather generator. Different colors represent simulations from two different parameter sets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

objective of getting a 'target' sample mean of $\bar{x} = 3$. Suppose that for one iteration the optimizer attempts to compare samples drawn from a simple Gaussian random generator ($X \sim N(\mu, \sigma)$) where the parameter μ is changed from 4.0 to 4.5, while holding σ at a constant value of 1. In the upper panel of Fig. 3, we show 50 random values drawn with each parameter set. For this set of random values, the sample mean from $X \sim N(4.0, 1)$ is 4.2 compared with the sample mean from $X \sim N(4.5, 1)$, which is 4.0. Therefore, the resulted sample mean from $N(4.0, 1)$ is actually further away from the target sample mean of $\bar{x} = 3$ compared with $N(4.5, 1)$, so that the search direction of the optimizer may be misled. Although this variance can be reduced with a larger sample size or a longer simulation period, it can never be completely eliminated.

To overcome this problem during optimization, the random number seed is held constant when producing the stochastic replicates. This ensures that any changes made to the parameters during the optimization process will lead the search in the desired direction. Using the same example, in the lower panel of Fig. 3 we show 50 samples drawn from both $X \sim N(4.0, 1)$ and $X \sim N(4.5, 1)$ with the same random seed used for each pair of samples, resulting in samples means of 3.9 and 4.4 respectively, thereby indicating that $N(4.0, 1)$ is better at producing a target sample mean of $\bar{x} = 3$. In this way, the stochastic generator is able to

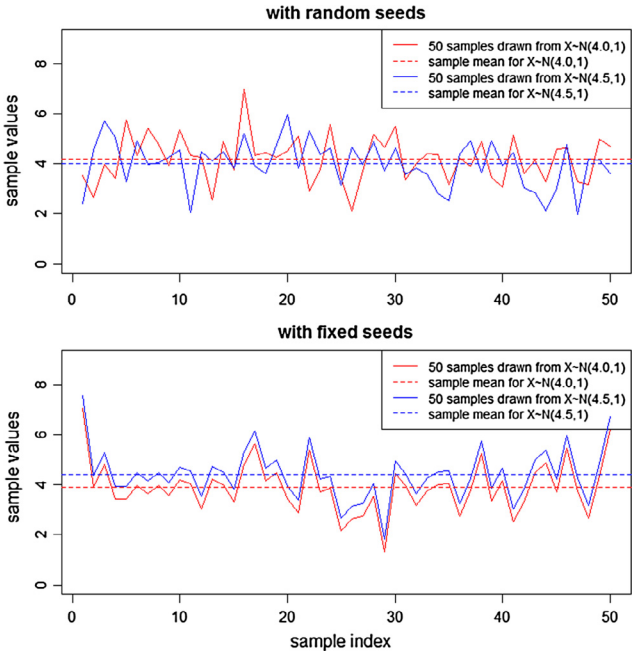

Fig. 3. Implications of using random seed (upper panel) and fixed seed (lower panel) for comparing two parameter sets: $X \sim N(4.0, 1)$ and $X \sim N(4.5, 1)$, for simulating a sample mean of 3 from a Gaussian distribution.

search through the correct directions on the parameter space to find parameters that converge toward the target rainfall attributes.

As discussed in Section 2.2, it is important to emphasize that the objective of the approach is to generate samples of hydrometeorological time series with specific levels for each attribute, rather than to identify the parameter sets that will produce those parameters in a population sense. Returning to the above example, the objective is to find a stochastic replicate with sample mean $\bar{x} = 3$, regardless of the values of the parameters (μ, σ) used to achieve this value. Consequently, once this goal has been met, the search can stop and the parameter values that were used to produce the stochastic time series corresponding to each target location can be discarded.

3. Case study

The proposed inverse approach is illustrated on rainfall data from a catchment in South Australia, using two stochastic rainfall generators: the Richardson model and the WGEN model. To provide a benchmark for the proposed inverse approach, its performance is also compared with that of a forward approach. The rainfall data, stochastic rainfall generators and the specific implementation of the forward and inverse approaches are described in this section.

3.1. Data

We used a rainfall time series from a gauge in the southern Mount Lofty Ranges close to Adelaide, South Australia, as a case study. The climate in this region is temperate, with most rainfall occurring in winter and spring (May to October). The mean annual rainfall was 913 mm for the study period from 1989 to 2004. The daily rainfall data over this period have been used to represent the baseline (historical) rainfall conditions.

We used four rainfall attributes as the dimensions of the exposure space, with definitions and baseline values provided in Table 1. These attributes represent key features of rainfall patterns; namely, the average daily rainfall (*PD*), the wet day frequency (*WD*), a measure of the rainfall intermittency (*CDD*) and a measure of extreme rainfall (*Pex99*). These attributes have been commonly used to assess the performance of rainfall generators (Chen and Brissette, 2014; Fowler et al., 2007; Hashmi et al., 2011; Kilsby et al., 2007; Semenov, 2007), and are also closely related to several of the indices used for the detection and attribution of climate change, as described by the Expert Team on Climate Change Detection and Indices (ETCCDI; Klein Tank et al., 2009).

For each rainfall attribute we defined a plausible range for sampling (which defined the range of each dimension within the exposure space) of between 50% and 150% of the corresponding historical value. These bounds were wider than would be expected from most climate change projections (e.g. CSIRO and Bureau of Meteorology, 2015; Stocker et al., 2013), to encompass a large range of climate projections (for example from climate models) in the exposure space.

Table 1
Definitions and annual average values of the four rainfall attributes that form the exposure space.

Rainfall attribute	Definition	Historical value
PD	Daily rainfall intensity averaged over all days	2.49 mm
WD	Average number of wet days	161 days
CDD	Average length of consecutive dry days	3.75 days
Pex99	The 99th percentile rainfall over wet days	37.6 mm

3.2. Stochastic rainfall generators

Two versions of the Richardson-type stochastic rainfall generator with different levels of complexity were used to generate the exposure space. We started with a simplified four-parameter model, which assumes uniform rainfall characteristics over the year. The advantage of this model is that it is possible to analytically derive the parameters that correspond to each target location in the exposure space. However, this simplified model uses a single value for each parameter throughout the year, and thus is unable to capture seasonal-scale variability. To highlight some practical issues with rainfall sampling, we then considered a more complex and widely used model—namely the WGEN (Richardson and Wright, 1984).

3.2.1. The four-parameter Richardson model

The simplified Richardson-type rainfall generator uses the following four parameters:

- The two parameters of the 1st order two-state Markov chain used for representing the transition probabilities of rainfall occurrence: p_{dd} (dry–dry probability) and p_{wd} (wet–dry probability), and
- The two parameters of a gamma distribution for representing the rainfall intensity on wet days: α (scale) and β (shape).

An approximate analytical expression relating two of the four output rainfall attributes (*PD* and *WD*) to the model parameters is given in Dubrovský et al. (2000) as:

$$PD = \frac{\alpha}{\beta} * \frac{WD}{365.25} \tag{2}$$

$$WD = 365.25 \times \frac{(1 - p_{dd})}{(1 - p_{dd} + p_{wd})} \tag{3}$$

These analytical expressions have been used when exploring the implications of random sampling issues on the inverse generation approach (Section 4.1.3).

3.2.2. The WGEN model

The WGEN model (Richardson and Wright, 1984) has the same structure as the simplified Richardson model, except that it uses a unique set of the four parameters for each month of the year, leading to a total of 48 parameters. This model has been used widely for climate impact studies, and is generally shown to capture most of the key features of daily rainfall series (Bastola et al., 2011; Katz, 2002; Kim et al., 2007).

Since the proposed inverse approach involves optimization of the parameter values, a search space with low dimension (i.e. consisting of a small number of parameters as decision variables) is desired. To reduce the size of the parameter space in the inverse approach, the number of decision variables to be considered was reduced from 48 to eight by fitting harmonic functions to describe the seasonal variations of each parameter (Prudhomme et al., 2013). The harmonic function takes the form of:

$$\beta(t) = \beta_0 + A\left(\cos\frac{2\pi}{T}(t - \Phi)\right) \tag{4}$$

where $\beta(t)$ represents one of the four parameters during month ($t = 1, \ldots, T$) with $T = 12$, β_0 represents the arithmetic mean of the parameter, A represents the amplitude and Φ corresponds to the month where the maximum occurs. It is worth mentioning that although parameter Φ can be varied as part of the optimization, the four-attribute exposure space in this case study was not designed to focus on shifts in rainfall seasonality (Section 3.1), so

that Φ was held constant at its historically optimal value. To determine the value of Φ, we obtained the monthly estimates of p_{dd}, p_{wd}, α and β (based on the method in Richardson (1981)) using the historical rainfall data, and fitted a harmonic function to each parameter (Fig. 4). The corresponding values of Φ were thus identified to be 2, 1, 8 and 1 (i.e., February, January, August and January) for the four parameters, respectively. As a result, the optimization was performed on the mean (β_0) and amplitude (A) of each of the four model parameters, leading to an eight dimensional search space.

3.3. Sampling approach

As illustrated in Fig. 1, application of the forward approach involves sampling the parameter space prior to using the stochastic model. Similarly, application of the inverse sampling approach involves the identification of target locations in the exposure space as the basis for optimization. One approach to sampling both the parameter space (in the forward approach) and exposure space (in the inverse approach) is to define a grid of evenly spaced points over the entire space. However, this can be inefficient, particularly for high-dimensional problems (for a large number of parameters/attributes in the exposure space in the forward/inverse approach). For example, if one wished to sample on a grid of width 10 for the parameter space of the four-parameter Richardson model, then it would be necessary to evaluate a total of $10^4 = 10{,}000$ separate parameter sets. This issue is particularly pertinent for the inverse approach, since optimization is required to find a parameter set that corresponds to each point in the exposure space. Therefore, to provide even coverage of the parameter or exposure space while keeping the sample sizes low, two structured sampling techniques have been employed, namely Latin hypercube sampling (LHS) and improved distributed hypercube sampling (IHS).

The objective of the analysis in this paper is to illustrate the inverse approach, by comparing its performance with the forward approach. Therefore for consistency, the objective of the sampling approach was to obtain 100 samples within the exposure space. For the forward approach, it is not known *a priori* whether a particular parameter set in the parameter space will yield a sample in our exposure space (i.e. within the plausible range of 50–150% for each rainfall attribute, as defined in Section 3.1), so that the number of samples that need to be drawn from the parameter space is not known. To determine the total number of samples in a computationally efficient manner, we used the Latin hypercube sampling (LHS) method, which allows starting with a small sample size and adding new samples while keeping the previously generated ones. The LHS method involves sampling M variables with a desired sample size N by dividing the range of each variable into N equally probable intervals. N samples are then drawn so that any interval for each variable is only sampled once (Stein, 1987). To add n new sample points, the existing design is re-divided into $(N + n)$ intervals; the N old samples are kept which occupy N intervals, and then n new samples are drawn to fill the remaining n intervals.

Unlike the LHS method, the IHS method (Manteufel, 2001; Beachkofski and Grandhi, 2002) requires that the number of samples be specified *a priori*, but ensures more even coverage of the sampling space. This latter feature is attractive when sparsely sampling potentially high-dimensional spaces, and is therefore recommended to determine the target locations in the exposure space for the inverse sampling approach. The IHS method is similar to the LHS method, with two additional objectives:

(1) The average minimum distance between sample points equals the optimal distance d_{opt}. That is, if the span of each output variable is normalized to 1 so that the entire sample space is a hypercube of volume 1, then each sample point should cover an equal hypervolume (with dimension of M) within the entire space. This gives the optimal distance between sample points, i.e. $d_{opt} = \sqrt[M]{\frac{1}{N_{Sample}}}$.

(2) The coefficient of variance (COV) of all minima between each pair of sample points is close to zero.

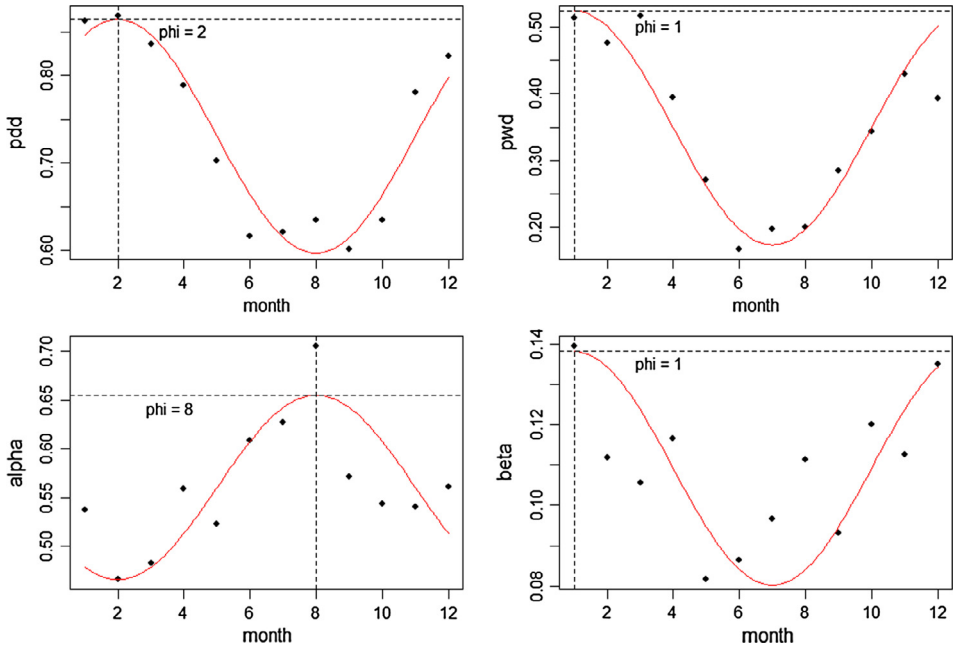

Fig. 4. The monthly variations of p_{dd}, p_{wd}, α and β for WGEN to determine the values of Φ, obtained from existing rainfall data, black dots represent individual parameter values estimated for each month, whereas red curves show the fitted harmonic functions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.4. Implementation of forward approach

As mentioned previously, the forward approach has been used to provide a benchmark against which the utility of the proposed inverse approach can be assessed. The approach involves the following steps:

(1) The parameter space is constructed by selecting appropriate ranges of the parameters for the stochastic generators;
(2) Parameter sets are drawn from the parameter space using a sampling strategy such as the LHS method described in Section 3.3;
(3) The sampled parameter sets are used to generate time series of the hydrometeorological variables of interest (in this case, rainfall); and
(4) The values of the attributes that define the exposure space are calculated for each of the generated hydrometeorological time series.

For the simple four-parameter Richardson model, the 'transition probability' parameters (p_{dd} and p_{wd}) both vary between 0 and 1. The two 'rainfall intensity' parameters (α and β) are for the gamma distribution and should be greater than 0 (Note that their values are mostly between 0 and 1 when calibrated to historical data; see Richardson and Wright, 1984). From a preliminary analysis for our case study, α and β values of 0.56 and 0.10 were obtained respectively, yielding rainfall time series with attributes that are close to the historical data. Therefore, although α and β do not physically have upper bounds and can take any value above 0, their ranges were set to be between 0 and 1 in the forward approach based on their historical values. The use of such a small range ensures that the parameter space surrounding the historical levels of the parameters is sufficiently sampled. For WGEN, the parameter ranges were defined in a similar way, so that the bounds of both the transition probability and rainfall intensity parameters were set to 0 and 1 for all months.

As mentioned previously, for both stochastic models, LHS was used to sample the parameter space. An initial Latin hypercube sample size of 100 parameter sets was used, and this was incremented until 100 rainfall time series were generated with attributes within the plausible bounds of the exposure space.

3.5. Implementation of proposed inverse approach

A general description of the inverse approach was provided in Section 2.2. The IHS method (Section 3.3) was used to determine the target locations for the optimization, which consist of 100 sets of combined levels of the four rainfall attributes that uniformly cover the exposure space.

For each target location, the best-fit parameter sets for both the four-parameter Richardson model and the WGEN model were identified using optimization. The shuffled complex evolution algorithm (Duan et al., 1993) was used as the optimization engine, due to its proven ability for solving complex optimization problems in hydrological studies (Gupta et al., 1999; Thyer et al., 1999; Wang et al., 2010). Based on the general formulation in Eq. (1), the objective function to be minimized for both stochastic models was:

$$F_{obj,i} = \sqrt{\begin{array}{l} \left\{\left[\frac{(PD_{s,i}-PD_{his})}{PD_{his}} - \frac{(\widehat{PD_i}-PD_{his})}{PD_{his}}\right]*100\right\}^2 + \left\{\left[\frac{(WD_{s,i}-WD_{his})}{WD_{his}} - \frac{(\widehat{WD_i}-WD_{his})}{WD_{his}}\right]*100\right\}^2 + \\ \left\{\left[\frac{(CDD_{s,i}-CDD_{his})}{CDD_{his}} - \frac{(\widehat{CDD_i}-CDD_{his})}{CDD_{his}}\right]*100\right\}^2 + \left\{\left[\frac{(Pex99_{s,i}-Pex99_{his})}{Pex99_{his}} - \frac{(\widehat{Pex99_i}-Pex99_{his})}{Pex99_{his}}\right]*100\right\}^2 \end{array}} \quad (5)$$

The constraints of the optimization consist of the plausible ranges of the parameters for both models. As mentioned in Section 3.4, the plausible range for the probability parameters (p_{dd}'s and p_{wd}'s) is between 0 and 1; for the intensity parameters (α's and β's), which do not have a physical upper limit, we defined the range to be between 0 to 10^4, which was wider than the range used for the forward approach (Section 3.4) to enable more extensive searching within the defined range.

For the WGEN model, since a harmonic function has been fitted to the monthly values of each of the probability and intensity parameters (Section 3.2.2), the actual decision variables for the optimization were the parameters of the harmonic functions (i.e. β_0 and A, which represent the mean and amplitude respectively, as in Eq. (4)). To ensure that the probability parameters were always within 0 and 1 while the intensity parameters were always within 0 and 10^4 during the optimization process, the values of the mean and amplitude for each of these parameters have been optimized sequentially. In the first step, the mean value of each parameter has been optimized with the amplitude kept as zero. Once the mean has been determined, a second optimization was conducted to estimate the amplitude. For example, if the mean of p_{dd} is found by the optimizer to be 0.3 in the first step, its amplitude must be constrained between 0 and 0.3 in the second step to avoid values of p_{dd} going beyond 0 and 1.

It should be noted that in determining the target locations, the IHS only checks the multi-dimensional uniformity of the overall distribution, without considering the physical interpretation for each individual target location. Therefore, it is important to ensure that each target location selected is physically plausible. For example, *PD* should always be less than *Pex99*, and *WD* should never exceed 365 days. For this study, these constraints were automatically satisfied because a relatively small plausible range of 50–150% was selected for each attribute. If the rainfall samples are required to show larger variances, it may be necessary to impose additional constraints in the optimization procedure to ensure the resultant samples remain physically plausible.

4. Results

4.1. The four-parameter Richardson model

4.1.1. Forward approach

The coverage of the exposure space obtained by applying the forward approach to the four-parameter Richardson model is shown in Fig. 5, which shows high variances in some rainfall attributes. In particular, the generated *PD*, *CDD* and *Pex99* can go up to 15,000%, 6000% and 80,000% of their corresponding historical values, respectively (Fig. 5a), which are well outside the bounds of the exposure space. The sampled *WD* has lower variance with values up to only 226% of the historical values (since a year contains a maximum of 365 or 366 days), however, these values are still above the upper limit of the exposure space of 150%. The high variance leads to low sampling efficiency – to obtain 100 sets of combined levels of rainfall attributes within our exposure space, a total of 7635 LHS samples of parameter sets had to be generated (i.e. 98.7% samples were unacceptable and discarded). All 7635 sets

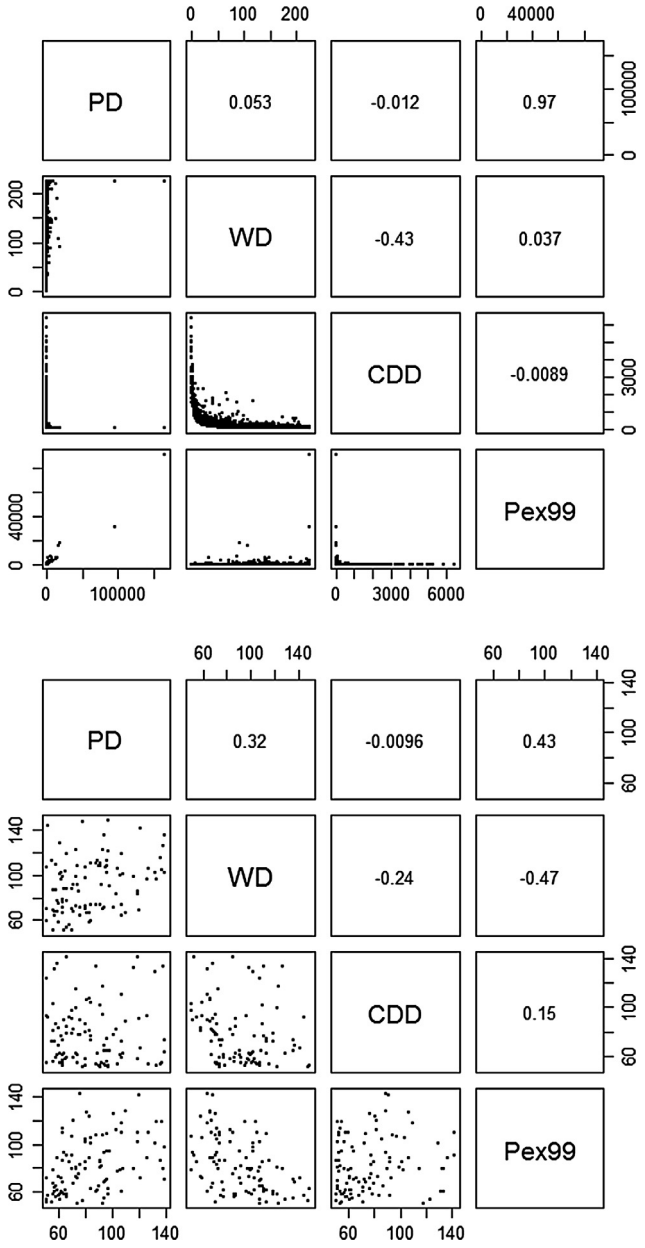

Fig. 5. The four rainfall attributes described as a percentage relative to the historical values produced by the four-parameter Richardson model, by drawing LHS samples from the parameter space, for (a) all the 7635 samples, and (b) only the 100 samples that fall in the plausible range of between 50% and 150% for each rainfall attribute. The upper-right triangle displays pairwise correlations.

of rainfall attributes are plotted in Fig. 5a, with the 100 plausible samples shown in Fig. 5b.

In addition to the issue of inefficient sampling, based on both a visual inspection of the coverage on the exposure space as well as consideration of the correlation coefficients, it is clear that the coverage of the exposure space is uneven (Fig. 5b). In particular, samples are clustered in small regions of the exposure space for each rainfall attribute, with other parts of the space receiving limited or no coverage. For example, the correlation between *PD* and *Pex99* is quite high, which results in better coverage over regions closer to the diagonal of the joint distribution of *PD* and *Pex99* than other regions.

The above problems with using the forward approach are most likely due to the non-linear translation from parameters to rainfall attributes through the stochastic generator, so that large variations in certain regions in parameters space result in small variations in exposure space and vice versa. This non-linearity will be further illustrated in the next section with the distribution of parameters identified through the inverse approach.

4.1.2. Inverse approach

Fig. 6a shows the 100 target locations of desired rainfall attributes that have been determined using the IHS approach (Section 3.3). As can be seen, the IHS approach generates samples that appear to be uniformly distributed across the exposure space, with even coverage across each attribute and low cross-correlations between attributes.

The final set of combined levels of attributes corresponding to each of the 100 stochastically generated rainfall time series obtained using the inverse approach is presented in Fig. 6b. As

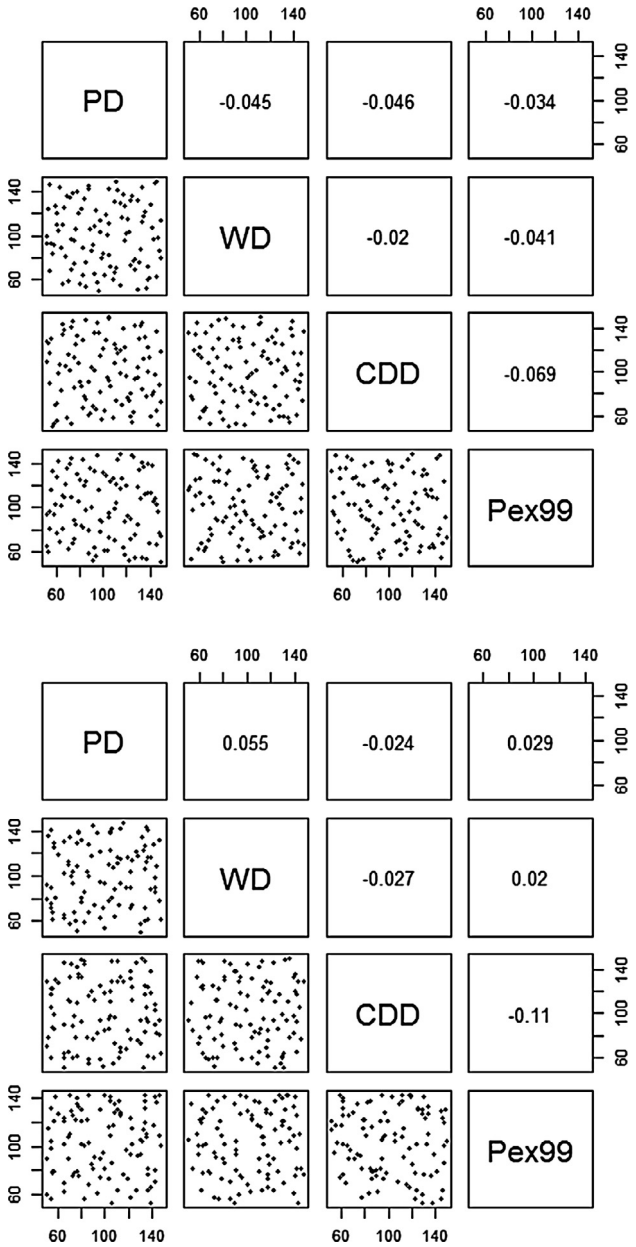

Fig. 6. The four rainfall attributes described as a percentage relative to the historical values produced by the four-parameter Richardson model for (a) 100 selected IHS sample locations and (b) 100 corresponding optimized locations. The upper-right triangle displays pairwise correlations.

can be seen, the optimization-based approach is effective in producing the desired levels of rainfall attributes (i.e. target locations), with all of the 100 samples falling within the bounds of the exposure space and with relatively even coverage of the exposure space (Fig. 6b). Therefore, the inverse approach delivers much better coverage of the exposure space than the forward approach (Fig. 5).

Fig. 7 shows the values of the 100 parameter sets for the four-parameter Richardson model, identified via application of the inverse approach, highlighting the non-linear mapping between parameter space and exposure space. This is best illustrated with the non-uniform distribution of the best-fit parameters, in contrast to the uniform distribution of the exposure space (Fig. 6). Furthermore, the parameters have considerably different ranges compared with the *a priori* [0,1] ranges that were specified for the forward approach. For example, the values of p_{dd} generally vary within a narrower range of 0.5–0.9, whereas values of α are as high as 10. Therefore, the ranges of [0,1] defined for the four parameters in the forward approach (as detailed in Section 3.4) can significantly limit the resultant coverage of the exposure space. This also reflects the high degree of non-linearity in the mapping between the parameter values and the exposure space, as a small change in the exposure space may result in a large shift in parameter space.

Interestingly, for the case study considered, although the inverse approach had the additional step of parameter optimization, the computational time required to obtain 100 samples was 32.6% shorter than for the forward approach. This is likely due to the large number of samples that were discarded in the forward approach.

4.1.3. Implications of random sampling on the inverse approach

In the above example, we fixed the random seed of the random number generator during the optimization process due to reasons discussed in Section 2.3. To illustrate the importance of this aspect of the optimization, we use the analytical expressions in Eqs. (2) and (3) to estimate the model parameters that will yield individual target locations from a grid consisting five evenly-spaced levels for each of *WD* and *PD* (50%, 75%, 100%, 125% and 150% of their historical values). These locations within the exposure space are given as green dots in Fig. 8. We then generated 100 stochastic replicates

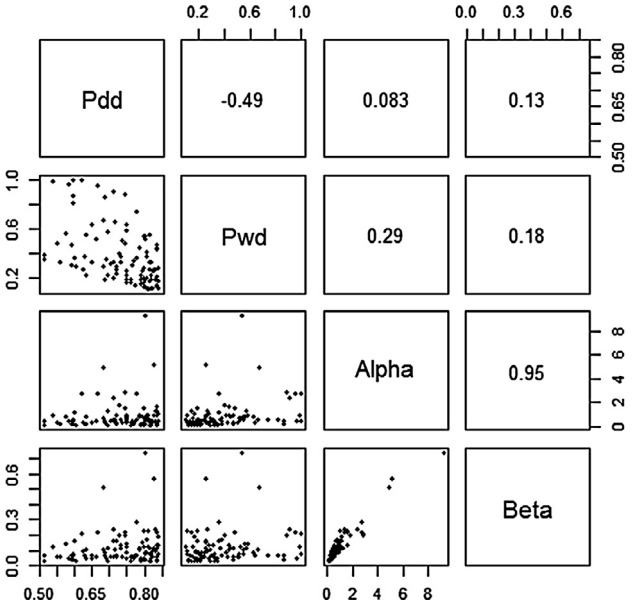

Fig. 7. Parameters corresponding to the exposure space in Fig. 6.

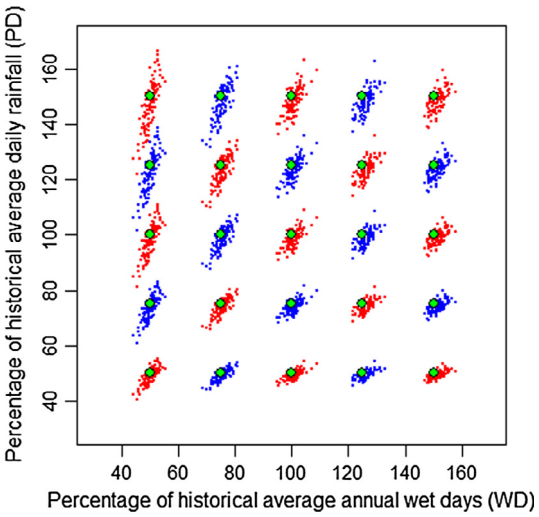

Fig. 8. Stochastic behavior of the four-parameter Richardson model in simulating *WD* and *PD*, with 100 replicates for the analytically-solved parameter set corresponding to each single target location. Two separate colors (i.e. red and blue) were used to differentiate between adjacent target locations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

from each of these parameter sets with different random seeds, which are shown as blue and red scatter about each of the target locations in Fig. 8.

The stochastic nature of the model is clear for all target locations. For each parameter set, the 100 replicates of *WD* vary up to ±10% around their target level, which is similar for all target levels of *WD* and *PD*. In contrast, the 100 replicates of *PD* are closer to the target level for smaller *PD* (e.g. up to ±10% around where the target level is 50%), while for larger *PD* target levels the spread among replicates increases substantially (e.g. up to ±40% around where the target level is 150%). Compared with the sampling resolution required in this study (shown in Fig. 6a), the variability in Fig. 8 is in fact much higher, which can adversely affect the capacity of the optimizer to find parameters that correspond to each target location, as discussed in Section 2.3.

4.2. The WGEN model

4.2.1. Forward approach

The coverage of the exposure space obtained by applying the forward approach to the WGEN model is shown in Fig. 9. Similar to the results for the four-parameter model (Section 4.1.1), the forward approach shows low efficiency: to obtain 100 sets of rainfall attributes within the range of the exposure space, 1453 LHS samples of WGEN parameter sets were required (Fig. 9a), which means that 93.1% of samples were discarded. With the 100 plausible sets in Fig. 9b, the coverage of the exposure space is poor, which is also evident through the high pairwise correlations (such as between *PD* and *Pex99* and between *WD* and *CDD*).

4.2.2. Inverse approach

To examine the performance of the inverse approach with WGEN, the 100 target locations which have been determined using the IHS approach (Section 3.3) are plotted in Fig. 10a. The final optimized set of attributes corresponding to each of the 100 stochastically generated rainfall time series is presented in Fig. 10b. The inverse approach is generally effective in evenly covering the exposure space and reproducing these target locations. In particular, this approach delivers much better coverage of the

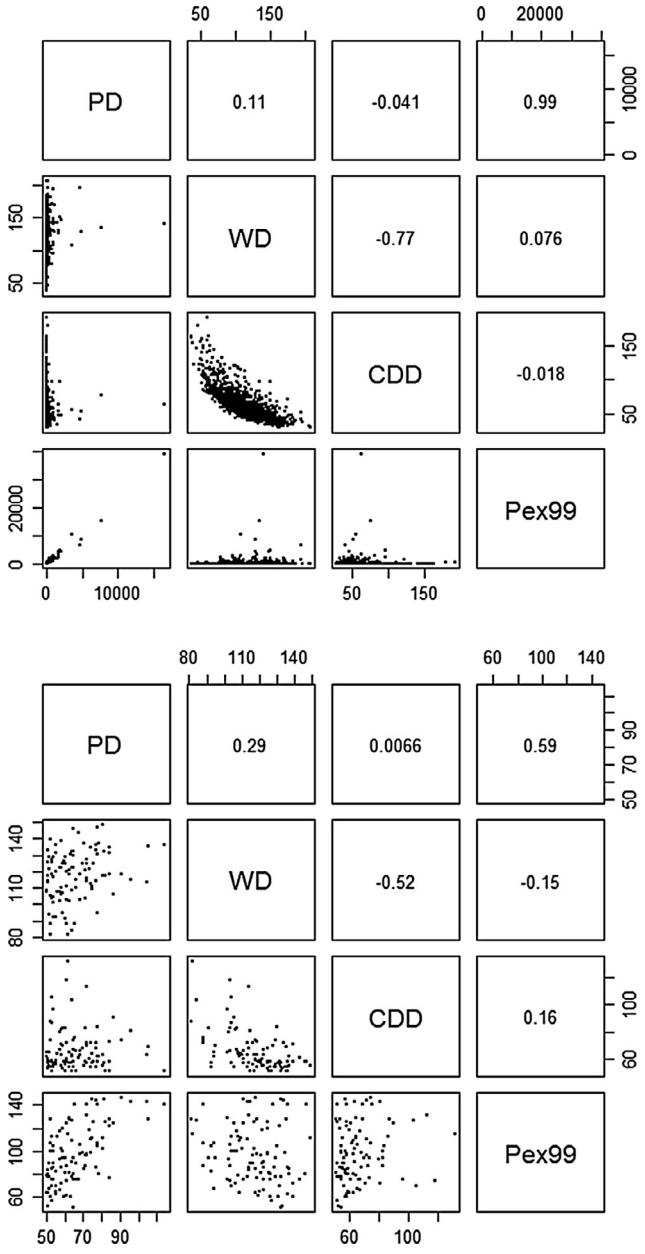

Fig. 9. The four rainfall attributes described as a percentage relative to the historical values produced by the WGEN model, by drawing LHS samples from the parameter space, for (a) all 1453 samples, and; (b) only the 100 samples that fall in the plausible range of between 50% and 150% for each rainfall attribute. The upper-right triangle display pairwise correlations.

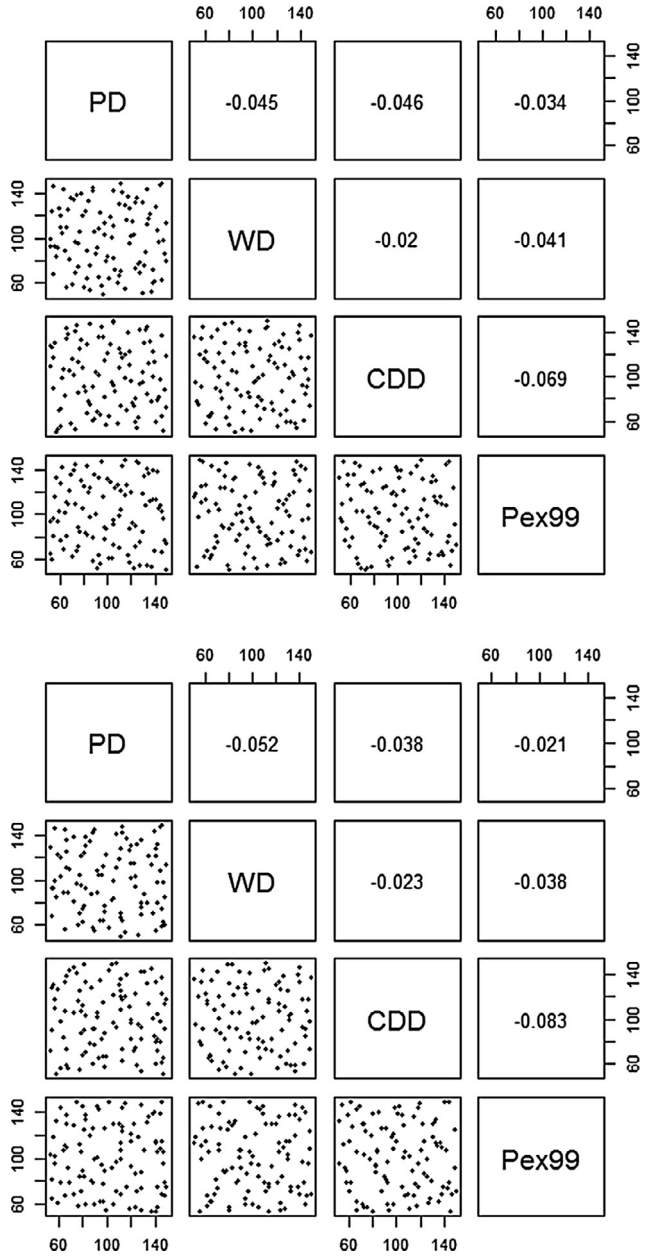

Fig. 10. The four rainfall attributes described as a percentage relative to the historical values produced by the WGEN model for (a) 100 selected IHS sample locations and (b) 100 corresponding optimized locations. The upper-right triangle displays pairwise correlations.

exposure space than the forward approach (Fig. 9) in the following aspects:

(1) All of the 100 samples are within the plausible output space defined in Table 1, suggesting effective control over the values of individual rainfall attributes; and
(2) The joint distribution of multiple rainfall attributes is much more uniform across the exposure space, and the pairwise correlations between different attributes are reduced.

As an alternative approach to assessing the relative uniformity of the sampling in the exposure space, the minimum distances between sample points in the exposure space are compared for both the forward (as orange dots in Fig. 11) and inverse approaches (as blue dots in Fig. 11). The results show that the inverse sampling approach produces a more uniform coverage, as the minima between sample points are closer to the optimal distance, $d_{opt} = 0.32$ (see Section 3.3). Furthermore, the coefficient of variance (*COV*) of these minimum distances is also much lower (i.e. 0.52 for the inverse approach compared with 2.90 for the forward approach).

It is worth noting that to obtain 100 sample points on the exposure space with the WGEN model, the overall execution time required for implementing the inverse approach is 73% longer than that for the forward approach. This is most likely due to the difficulty in solving optimization problems with a larger number of parameters, as a result of the larger search space that has to be explored. However, the inverse approach ensures uniform coverage of the exposure space with the desired resolution, which is the key objective for constructing the exposure space. In contrast,

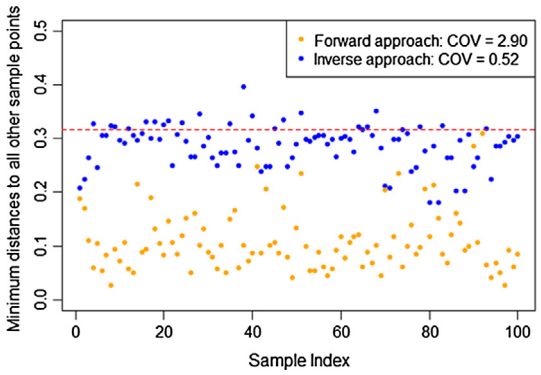

Fig. 11. Minimum distances between sample points obtained from the forward and inverse approaches with the WGEN model. The dashed line shows the optimal distance with 100 plausible samples (d_{opt} = 0.32).

the forward approach failed to obtain such coverage. Therefore, although associated with a higher computational expense, the proposed inverse approach is the only way of achieving the desired coverage of the exposure space.

5. Discussion

This study presented a framework for sampling various rainfall conditions to construct an exposure space for scenario-neutral climate impact assessments. Here, we discuss some practical considerations, as well as possible future adaptations of the framework.

5.1. Design of exposure space to represent more complex potential climate changes

The four rainfall attributes considered in the exposure space for this study (i.e. *PD*, *WD*, *CDD* and *Pex99*) are good descriptors of a range of changes of annual precipitation characteristics. However, there is also a range of other rainfall attributes that might be important when considering the impact of climate change, such as changes at seasonal or interannual timescales (e.g. Christensen et al., 2007; Johnson et al., 2011; Kwon et al., 2009). Furthermore, potential future variations in other climatic features, such as temperature, solar radiation and evapotranspiration, may also have a substantial impact on water resources (for examples see Chiew and McMahon, 2002a; Prudhomme and Williamson, 2013).

The inverse approach presented here is sufficiently flexible to cater to all attributes in the exposure space that are of interest (provided they can be generated with an appropriate stochastic generator), although this comes at the expense of additional computational cost. Considering the trade-off between the flexibility of producing different climate patterns and computational effort, it is important to identify key hydrometeorological variables of interest, as well as their attributes, based on an understanding of the behavior of the system being analyzed.

Depending on the specific hydrometeorological variables involved, the format of the objective function may require modification from Eq. (1), which was designed assuming multiplicative perturbations to attributes (e.g. changes expressed as a percentage of the historical value). For example, temperature changes are typically represented in an additive way (e.g. increases in temperature by degrees Celsius; for examples see Chiew and McMahon, 2002a; Kingston et al., 2009), and this would require an adjustment to the objective function in Eq. (1).

In this study, the boundary of the exposure space was set at 50–150% of the historical values of each attribute, which is sufficiently wide to incorporate a large number of possible changes in each of the rainfall attributes, while also using the same percentage changes across attributes to facilitate illustration. However, this framework can be easily adapted to incorporate tailored bounds for the exposure space, which should be carefully selected to suit the case study under consideration. In particular, if the bounds deviate too far from present conditions, a significant portion of samples will be unrealistic, even when extreme climate change impacts are considered. Conversely, if the bounds are too narrow, system response to some plausible climatic changes might not be considered (Whateley et al., 2014). Multiple sources of information could be considered in selecting these bounds, including GCM-based climate projections (e.g. Collins et al., 2013) of possible future climatic changes, and additional lines of evidence on possible changes to key variables, such as from long-term paleoclimatology reconstructions (e.g. Ault et al., 2014; Hansen and Sato, 2012; Ho et al., 2015). In addition, it is worth specifying an exposure space with bounds that are wider than the range suggested from all currently available sources of information, so that additional climate change projections can be included in the analysis as they are developed (Steinschneider and Brown, 2013).

Finally, when determining the target locations consisting of different hydrometeorological attributes on the exposure space, it is desirable to ensure the physical realism of each individual location so that corresponding time series can be obtained with the aid of stochastic weather generators. This requires not only ensuring that the target levels of individual attribute are realistic (such as the constraints for the levels of *WD*, as discussed in Section 3.5), but also maintaining physically plausible relationships among multiple attributes (for example, a target location cannot consist a *WD* value of 100 days with a *CDD* value of 300 days, because this combination means that the annual average wet day is 100 days while the annual average dry spell length is 300 days, which is physically unrealistic).

5.2. Stochastic generation of the exposure space

In this study, we used a sample size of 100 to represent different levels of changes for each individual attribute considered in the exposure space, with fixing the random seed across replicates to facilitate improved convergence during the optimization process. In this way, however, there is likely to be limited variability in between time series corresponding to different points on the exposure space, except for variations related to the target statistics.

This issue can be addressed in at least three ways:

1. Increase the sample size, and thus the coverage resolution in the exposure space. Increasing the exposure space resolution is likely to be particularly useful when the number of attributes increases, as this will lead to a corresponding increase in the dimensionality of the exposure space.
2. The length of each sample can be increased. Currently, the length is equal to the length of the historical data series (i.e. 15 years). However, it would be trivial to allow the simulation to run for longer periods of time to obtain greater stochastic variation. This will require the same number of optimized parameter sets, although because of the use of the same initial seed, there will still be significant similarities between individual samples.
3. The procedure can be repeated multiple times with different random seeds for each iteration, thereby generating multiple replicates. This would substantially increase the level of stochastic variability, although at the expense of additional computational time.

Table A.1
Definition of terminology used in this study.

Terminology	Definition
Exposure space	A large range of plausible hydrometeorological conditions that a system may confront in the future
Hydrometeorological attributes	The statistics calculated from the time-series of hydrometeorological variables (e.g. rainfall, temperature and potential evapotranspiration) which describe their average conditions, variations, extremes, as well as other features. For example, temperature is considered as a hydrometeorological variable, with potential attributes including annual average temperature, annual maximum temperature and seasonal average temperature, etc.
Low-dimensional exposure space	An exposure space that consists of a small number of hydrometeorological variables and their attributes
High-dimensional exposure space	An exposure space that consists of a large number of hydrometeorological variables and their attributes
Forward approach	The approach of perturbing the parameters of stochastic generators over some pre-defined 'parameter space' to yield an exposure space
Historical scaling	The approach to obtain a desired exposure by applying additive and/or multiplicative scaling factors directly to the historical time series of hydrometeorological variables (as in Kay et al., 2014; Prudhomme et al., 2010; Prudhomme and Williamson, 2013)
Forward-plus-scaling approach	The approach that combines the use of the forward and scaling techniques to generate an exposure space. In this study, this specifically refers to the approach illustrated in Steinschneider and Brown (2013). In this approach, the parameters of a stochastic generator (including Markov chain transition probabilities and the autoregressive model for low-frequency variability) are first perturbed to obtain stochastic sequences without changing the historical rainfall intensity. The wet-day rainfall intensity in the stochastic sequences is subsequently quantile-mapped to yield a set of daily rainfall series with desired levels of rainfall attributes
Inverse approach	The approach aims to generate an exposure space by first selecting the desired values of the attributes of interest in the exposure space, followed by an optimization step to identify the stochastic generator parameters that produce stochastic sequences with these attributes
Target level	A desired level for a single meteorological attribute that is included in the exposure space
Target location	A desired combination of the target levels for each of the hydrometeorological attributes that are included in the exposure space
Best-fit parameters for the stochastic generator	The parameter set for the stochastic generator that produces hydrometeorological time series with the levels of attributes corresponding to a particular target location

The appropriate sample size, length of each sample and number of replicates are likely to depend on the case study considered, as well as available computational resources.

5.3. Equi-finality in the optimization process

When using optimization to search for best-fit solutions, equi-finality issues are likely to arise (i.e. multiple solutions leading to the same results in the objective function, therefore they are not distinguishable during optimization, see Beven and Freer, 2001). This problem is further complicated within the proposed inverse approach, as the values of the objective function for optimization are based on results from stochastic models.

Equi-finality issues are likely to be greatest for low-dimensional exposure spaces, since higher-dimensional exposure spaces add constraints to the parameter space. For example, the chance that two contrasting combinations of P_{dd} and P_{wd} lead to the same combination of *WD* and *CDD* is much lower compared to that resulting in the same level of *WD* in isolation. Thus, increasing the number of attributes considered could have the additional advantage of reducing the number of feasible parameter sets to be considered.

It is worth noting that although equi-finality is likely to occur when the proposed inverse approach is implemented, the aim of the approach is to identify time series of outputs from the stochastic generator that result in desired values of the attributes included in the exposure space, and not to the identification of the resulting parameters in the stochastic generator, as discussed in Section 2.3. However, when different parameter sets lead to the same combination of attributes on the exposure space, the different time series of hydrometeorological variables which they produce can consist of varying degrees of physical realism. Therefore, checking the physical realism of the generated time series can potentially help to eliminate unrealistic parameter sets and thus resolve any equi-finality issues.

5.4. Computational efficiency and execution time

In our particular implementation of the proposed inverse approach, the computational time required to produce 100 evenly distributed samples is around eight hours using an Intel Xeon E3 (2.60 GHz, 8 Cores) processor with 32 GB RAM for both the four-parameter model and the WGEN, suggesting a relatively high computational demand. However, in general, the computational effort required is dependent on a number of practical specifications, including the operating system, programming language and algorithm used. As the key focus of this study is to introduce and illustrate a new method, the above-mentioned specifications have not been optimized for computational efficiency. We have used the R package *lhs* (Carnell, 2012) for obtaining samples over the exposure space with the IHS and LHS methods, together with the shuffled complex evolution algorithm embedded in the R package *hydromad* (Andrews and Guillaume, 2013) for solving the optimal parameter values for the stochastic generator. We have also developed our own R-scripts to execute the four-parameter Richardson and the WGEN models. It is expected that an improved integration of these different modeling components with the aid of other programming languages, such as Fortran or C++, will further increase computational efficiency.

6. Conclusions

Generation of exposure spaces for scenario-neutral climate impact assessments should consider a range of potential variations in relevant hydrometeorological variables, including shifts in the average, intermittency, variability and extremes. The 'exposure space' describes the range of conditions of interest that a system may be exposed to under a future climate, and this paper presents and demonstrates an inverse approach to stochastically generating hydrometeorological time series to uniformly cover this exposure space.

The utility of the proposed inverse approach is benchmarked against a forward approach for rainfall generation for a South Australian catchment, using two Richardson-type stochastic rainfall generators of varying complexity. The results highlight the highly non-linear translation from parameter space to exposure space, and thus the need for the proposed inverse approach in order to

obtain a relatively uniform coverage of the exposure space. For both models, the inverse approach demonstrates better control of the sampling range, with 100% of samples falling within the exposure space. Furthermore, the uniformity of the coverage of the four-dimensional exposure space is substantially improved.

Several potential adaptations for future implementations of the framework have been discussed, including: (1) the design of the exposure space to represent more complex changes in climate; (2) improvements to the way that stochastic samples in the exposure space are generated; (3) ways of reducing the effects of equifinality during the optimization process; and (4) methods for increasing computational efficiency. The flexibility of the proposed inverse approach enables consideration of all climate attributes of interest at the desired resolution, thereby expanding the applicability of the scenario-neutral approach to evaluating a water resource system's sensitivity to a wide range of plausible changes in climate.

Acknowledgements

The authors wish to thank Christel Prudhomme and an anonymous reviewer for their thoughtful comments on the manuscript. Seth Westra's time was supported by Australian Research Council Discovery project DP150100411.

Appendix A

See Table A.1.

References

Ajami, N.K., Duan, Q., Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour. Res. 43 (1), W01403. http://dx.doi.org/10.1029/2005WR004745.

Alexander, L. et al., 2006. Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res.: Atmos. 111 (D5).

Allen, R., Pruitt, W., 1986. Rational use of the FAO Blaney–Criddle Formula. J. Irrig. Drain. Eng. 112 (2), 139–155. http://dx.doi.org/10.1061/(ASCE)0733-9437(1986) 112:2(139).

Andrews, F., Guillaume, J., 2013. Hydromad: Hydrological Model Assessment and Development. R package version 0.9-18.

Ault, T.R., Cole, J.E., Overpeck, J.T., Pederson, G.T., Meko, D.M., 2014. Assessing the risk of persistent drought using climate model simulations and paleoclimate data. J. Clim. 27 (20), 7529–7549. http://dx.doi.org/10.1175/JCLI-D-12-00282.1.

Bastola, S., Murphy, C., Sweeney, J., 2011. The sensitivity of fluvial flood risk in Irish catchments to the range of IPCC AR4 climate change scenarios. Sci. Total Environ. 409 (24), 5403–5415. http://dx.doi.org/10.1016/j.scitotenv.2011.08.042.

Beachkofski, B.K., Grandhi, R., 2002. Improved distributed hypercube sampling.

Beven, K., 1987. Towards the use of catchment geomorphology in flood frequency predictions. Earth Surf. Proc. Land. 12 (1), 69–82.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249 (1–4), 11–29. http://dx.doi.org/10.1016/S0022-1694(01)00421-8.

Boughton, W., Droop, O., 2003. Continuous simulation for design flood estimation—a review. Environ. Modell. Softw. 18 (4), 309–318. http://dx.doi.org/10.1016/S1364-8152(03)00004-5.

Brown, C., Ghile, Y., Laverty, M., Li, K., 2012. Decision scaling: linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resour. Res. 48 (9), W09537. http://dx.doi.org/10.1029/2011WR011212.

Brown, C., Werick, W., Leger, W., Fay, D., 2011. A decision-analytic approach to managing climate risks: application to the upper great lakes1. JAWRA J. Am. Water Resour. Assoc. 47 (3), 524–534. http://dx.doi.org/10.1111/j.1752-1688.2011.00552.x.

Brown, C., Wilby, R.L., 2012. An alternate approach to assessing climate risks. Eos, Trans. Am. Geophys. Union 93 (41), 401–402.

Carnell, R., 2012. lhs: Latin Hypercube Samples. R package version 0.10.

Chen, J., Brissette, F.P., 2014. Comparison of five stochastic weather generators in simulating daily precipitation and temperature for the Loess Plateau of China. Int. J. Climatol. 34 (10), 3089–3105.

Chiew, F.H., McMahon, T.A., 2002a. Modelling the impacts of climate change on Australian streamflow. Hydrol. Process. 16 (6), 1235–1245.

Chiew, F.H.S., McMahon, T.A., 2002b. Modelling the impacts of climate change on Australian streamflow. Hydrol. Process. 16 (6), 1235–1245. http://dx.doi.org/10.1002/hyp.1059.

Christensen, J.H. et al., 2007. Regional climate projections. Climate Change, 2007: The Physical Science Basis. Contribution of Working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. University Press, Cambridge, pp. 847–940, Chapter 11.

Clark, M.P., Slater, A.G., 2006. Probabilistic quantitative precipitation estimation in complex terrain. J. Hydrometeorol. 7 (1), 3–22. http://dx.doi.org/10.1175/JHM474.1.

Collins, M. et al., 2013. Long-term climate change: projections, commitments and irreversibility. In: Stocker, T.F. et al. (Eds.), Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1029–1136. http://dx.doi.org/10.1017/CBO9781107415324.024.

CSIRO and Bureau of Meteorology, 2015. Climate Change in Australia Information for Australia's Natural Resource Management Regions: Technical Report. CSIRO and Bureau of Meteorology, Australia.

Dessai, S., Hulme, M., 2004. Does climate adaptation policy need probabilities? Clim. Policy 4 (2), 107–128.

Duan, Q., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76 (3), 501–521.

Dubrovský, M., Žalud, Z., Šťastná, M., 2000. Sensitivity of CERES-Maize yields to statistical structure of daily weather series. Clim. Change 46 (4), 447–472.

Fowler, H., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27 (12), 1547–1578.

Frost, A., 2004. Stochastic generation of point rainfall data at subdaily timescales: a comparison of DRIP and NSRP.

Furrer, E.M., Katz, R.W., 2008. Improving the simulation of extreme precipitation events by stochastic weather generators. Water Resour. Res. 44 (12).

Gao, L. et al., 2016. Robust global sensitivity analysis under deep uncertainty via scenario analysis. Environ. Modell. Softw. 76, 154–166. http://dx.doi.org/10.1016/j.envsoft.2015.11.001.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. J. Hydrol. Eng.

Halton, J.H., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Num. Math. 2 (1), 84–90. http://dx.doi.org/10.1007/BF01386213.

Hammersley, J.M., 1960. Monte Carlo methods for solving multivariable problems. Ann. N. Y. Acad. Sci. 86 (3), 844–874. http://dx.doi.org/10.1111/j.1749-6632.1960.tb42846.x.

Hansen, J.E., Sato, M., 2012. Paleoclimate implications for human-made climate change. In: Berger, A., Mesinger, F., Sijacki, D. (Eds.), Climate Change: Inferences from Paleoclimate and Regional Aspects. Springer Vienna, Vienna, pp. 21–47. http://dx.doi.org/10.1007/978-3-7091-0973-1_2.

Hashmi, M.Z., Shamseldin, A.Y., Melville, B.W., 2011. Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. Stoch Environ. Res. Ris. Assess. 25 (4), 475–484.

Ho, M., Kiem, A.S., Verdon-Kidd, D.C., 2015. A paleoclimate rainfall reconstruction in the Murray-Darling Basin (MDB), Australia: 1. Evaluation of different paleoclimate archives, rainfall networks, and reconstruction techniques. Water Resour. Res. 51 (10), 8362–8379. http://dx.doi.org/10.1002/2015WR017058.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press.

Johnson, F., Westra, S., Sharma, A., Pitman, A.J., 2011. An assessment of GCM skill in simulating persistence across multiple time scales. J. Clim. 24 (14), 3609–3623. http://dx.doi.org/10.1175/2011JCLI3732.1.

Jones, P., Harpham, C., Goodess, C., Kilsby, C., 2011. Perturbing a weather generator using change factors derived from regional climate model simulations. Nonlinear Process. Geophyss. 18 (4), 503–511.

Jones, P.G., Thornton, P.K., 1993. A rainfall generator for agricultural applications in the tropics. Agric. For. Meteorol. 63 (1), 1–19.

Jones, R., Page, C., 2001. Assessing the risk of climate change on the water resources of the Macquarie River Catchment. Integr. Models Nat. Resour. Manage. Across Discip., Issues Scales, 673–678.

Katz, R.W., 2002. Techniques for estimating uncertainty in climate change scenarios and impact studies. Clim. Res. 20 (2), 167–185.

Kay, A.L., Crooks, S.M., Reynard, N.S., 2014. Using response surfaces to estimate impacts of climate change on flood peaks: assessment of uncertainty. Hydrol. Process. 28 (20), 5273–5287. http://dx.doi.org/10.1002/hyp.10000.

Kay, A.L., Jones, R.G., 2012. Comparison of the use of alternative UKCP09 products for modelling the impacts of climate change on flood frequency. Clim. Change 114 (2), 211–230. http://dx.doi.org/10.1007/s10584-011-0395-z.

Kilsby, C. et al., 2007. A daily weather generator for use in climate change studies. Environ. Modell. Softw. 22 (12), 1705–1719.

Kim, B., Kim, H., Seoh, B., Kim, N., 2007. Impact of climate change on water resources in Yongdam Dam Basin, Korea. Stoch. Environ. Res. Ris. Assess. 21 (4), 355–373. http://dx.doi.org/10.1007/s00477-006-0070-5.

Kingston, D.G., Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W., 2009. Uncertainty in the estimation of potential evapotranspiration under climate

change. Geophys. Res. Lett. 36 (20), L20403. http://dx.doi.org/10.1029/2009GL040267.
Klein Tank, A.M.G., Zwiers, F.W., Zhang, X., 2009. Guidelines on Analysis of Extremes in a Changing Climate in Support of Informed Decisions for Adaptation. World Meteorological Organisation.
Kwon, H.-H., Lall, U., Obeysekera, J., 2009. Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida. Stoch. Environ. Res. Ris. Assess. 23 (7), 879–896.
Langousis, A., Kaleris, V., 2014. Statistical framework to simulate daily rainfall series conditional on upper-air predictor variables. Water Resour. Res. 50 (5), 3907–3932. http://dx.doi.org/10.1002/2013WR014936.
Langousis, A., Mamalakis, A., Deidda, R., Marrocu, M., 2015. Assessing the relative effectiveness of statistical downscaling and distribution mapping in reproducing rainfall statistics based on climate model results. Water Resour. Res. http://dx.doi.org/10.1002/2015WR017556, n/a-n/a.
Maier, H. et al., 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. Environ. Modell. Softw. 62, 271–299.
Manteufel, R.D., 2001. Distributed hypercube sampling algorithm, Third AIAA non-deterministic approaches forum paper AIAA-2001-1673. In: 42nd structures, structural dynamics, and materials conference.
Meselhe, E., Habib, E., Oche, O., Gautam, S., 2009. Sensitivity of conceptual and physically based hydrologic models to temporal and spatial rainfall sampling. J. Hydrol. Eng. 14 (7), 711–720. http://dx.doi.org/10.1061/(ASCE)1084-0699(2009) 14:7(711).
Moody, P., Brown, C., 2013. Robustness indicators for evaluation under climate change: application to the upper Great Lakes. Water Resour. Res. 49 (6), 3576–3588.
Nazemi, A., Wheater, H.S., 2014. Assessing the vulnerability of water supply to changing streamflow conditions. Eos, Trans. Am. Geophys. Union 95 (32), 288-288.
Nazemi, A., Wheater, H.S., Chun, K.P., Elshorbagy, A., 2013. A stochastic reconstruction framework for analysis of water resource system vulnerability to climate-induced changes in river flow regime. Water Resour. Res. 49 (1), 291–305. http://dx.doi.org/10.1029/2012WR012755.
Paton, F., Maier, H., Dandy, G., 2013. Relative magnitudes of sources of uncertainty in assessing climate change impacts on water supply security for the southern Adelaide water supply system. Water Resour. Res. 49 (3), 1643–1667.
Poff, N.L. et al., 2015. Sustainable water management under future uncertainty with eco-engineering decision scaling. Nat. Clim. Change. http://dx.doi.org/10.1038/nclimate2765, advance online publication.
Prudhomme, C., Crooks, S., Kay, A., Reynard, N., 2013. Climate change and river flooding: part 1 classifying the sensitivity of British catchments. Clim. Change 119 (3–4), 933–948. http://dx.doi.org/10.1007/s10584-013-0748-x.
Prudhomme, C., Wilby, R.L., Crooks, S., Kay, A.L., Reynard, N.S., 2010. Scenario-neutral approach to climate change impact studies: application to flood risk. J. Hydrol. 390 (3–4), 198–209. http://dx.doi.org/10.1016/j.jhydrol.2010.06.043.
Prudhomme, C., Williamson, J., 2013. Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. Hydrol. Earth Syst. Sci. 17 (4), 1365–1377.
Racsko, P., Szeidl, L., Semenov, M., 1991. A serial approach to local stochastic weather models. Ecol. Model. 57 (1), 27–41.
Rajah, K. et al., 2014. Changes to the temporal distribution of daily precipitation. Geophys. Res. Lett. 41 (24), 8887–8894.
Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res. 17 (1), 182–190.
Richardson, C.W., Wright, D.A., 1984. WGEN: A Model for Generating Daily Weather Variables. US Department of Agriculture, Agricultural Research Service Washington, DC, USA.
Semenov, M.A., 2007. Simulation of extreme weather events by a stochastic weather generator. Clim. Res. 35 (3), 203.
Semenov, M.A., Brooks, R.J., 1999. Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. Clim. Res. 11 (2), 137–148.
Singh, R., Wagener, T., Crane, R., Mann, M.E., Ning, L., 2014. A vulnerability driven approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: application to a watershed in Pennsylvania, USA. Water Resour. Res. 50 (4), 3409–3427. http://dx.doi.org/10.1002/2013WR014988.
Stein, M., 1987. Large sample properties of simulations using Latin hypercube sampling. Technometrics 29 (2), 143–151.
Steinschneider, S., Brown, C., 2013. A semiparametric multivariate, multi-site weather generator with low-frequency variability for use in climate risk assessments. Water Resour. Res. http://dx.doi.org/10.1002/wrcr.20528, n/a-n/a.
Stocker, T.F. et al., 2013. Climate change 2013: The physical science basis. Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5). Cambridge Univ Press, New York.
Thyer, M., Kuczera, G., Bates, B.C., 1999. Probabilistic optimization for conceptual rainfall–runoff models: a comparison of the shuffled complex evolution and simulated annealing algorithms. Water Resour. Res. 35 (3), 767–773.
Vano, J.A., Kim, J.B., Rupp, D.E., Mote, P.W., 2015. Selecting climate change scenarios using impact-relevant sensitivities. Geophys. Res. Lett. 42 (13), 5516–5525.
Wang, Y.C., Yu, P.S., Yang, T.C., 2010. Comparison of genetic algorithms and shuffled complex evolution approach for calibrating distributed rainfall–runoff model. Hydrol. Process. 24 (8), 1015–1026.
Westra, S., Alexander, L.V., Zwiers, F.W., 2013. Global increasing trends in annual maximum daily precipitation. J. Clim. 26 (11), 3904–3918.
Westra, S. et al., 2014. Future changes to the intensity and frequency of short-duration extreme rainfall. Rev. Geophys. 52 (3), 522–555. http://dx.doi.org/10.1002/2014RG000464.
Whateley, S., Steinschneider, S., Brown, C., 2014. A climate change range-based method for estimating robustness for water resources supply. Water Resour. Res. 50 (11), 8944–8961.
Wilby, R.L., Dawson, C.W., Murphy, C., O'Connor, P., Hawkins, E., 2014. The Statistical DownScaling Model – Decision Centric (SDSM–DC): conceptual basis and applications. Clim. Res. 61 (3), 259–276.
Yates, D.N., Miller, K.A., Wilby, R.L., Kaatz, L., 2015. Decision-centric adaptation appraisal for water management across Colorado's Continental Divide. Clim. Risk Manage. 10, 35–50. http://dx.doi.org/10.1016/j.crm.2015.06.001.